

automated assistant **110** of FIG. 1 may be implemented on the client device **406** and/or on one or more computing devices that are in network communication with the client device **406**. Accordingly, for ease in explanation the automated assistant **110** is referenced in description of FIGS. 4, 5, and 6.

In FIG. 4, the user provides spoken input **480A** of “Buy 2 adult tickets for Movie A for tonight”. Voice input corresponding to the spoken input is generated by the device **606** and provided to the automated assistant **110** (e.g., as streaming voice input). The automated assistant **110** may convert the voice input to text, and determine that the text indicates an intended action of “buy movie ticket”. The automated assistant **110** may further determine that the text indicates a value of “2” for the parameter of “number of adult tickets”, and a value of “Movie A” (or associated identifier) for the parameter of “movie name”.

The automated assistant **110** may further determine an additional mandatory parameter of “theater location” that is defined for the “buy movie ticket” intended action is not specified by the spoken input **480A**. In response, the automated assistant **110** may generate and provide the prompt **482A** “what theater” to solicit a value for the mandatory parameter.

In response to the responsive output **482A**, the user provides further spoken input **480B** of “Baxter Avenue”. Voice input corresponding to the spoken input is generated by the device **606** and provided to the automated assistant **110**. The automated assistant **110** may convert the voice input to text, and based on the text determine “Baxter Avenue Theater” (or associated identifier) as a value for the mandatory parameter of “theatre location”.

The automated assistant **110** may further determine that two 3P agents (“3P1” and “3P2”) are able to perform the “buy movie ticket” intended action with the values of “2”, “Movie A”, and “Baxter Avenue Theatre”. In some implementations, the automated assistant **110** may select those from a larger group of 3P agents that can perform the “buy movie ticket” intended action based on the selected 3P agents being able to perform the intended action with the determined values, whereas others of the larger group cannot.

The automated assistant **110** further generates and provides the prompt **482B** that identifies those selected 3P agents, as well as values for an additional unspecified parameter (price) for each of the 3P agents.

In response to the responsive output **482B**, the user provides further spoken input **480C** of “I’ll use 3P2”. Voice input corresponding to the spoken input is generated by the device **606** and provided to the automated assistant **110**. The automated assistant **110** may convert the voice input to text, and based on the text determine 3P agent 3P2 is selected. The automated assistant **110** may then send an invocation request to 3P agent 3P2, along with determined values for parameters.

The 3P agent 3P2 then effectively or actually takes over (i.e., steers) the dialog and prompt **482C** is provided based on responsive content generated by 3P agent 3P2. Prompt **482C** may be provided directly by 3P agent 3P2 and/or via the automated assistant **110**.

In response to the responsive output **482B**, the user provides further spoken input **480C** of “Buy the 7:00 tickets”. Voice input corresponding to the spoken input is generated by the device **606** and provided to the automated assistant **110** (which then forwards it and/or a conversion to the 3P agent) and/or to the 3P agent 3P2 directly.

The 3P agent 3P2 generates further responsive content based on the spoken input **480D** (and/or a conversion thereof) and responsive output **482D** is provided based on that responsive content. Steering of the dialog is then returned to the automated assistant **110**.

The user then provides spoken input **480E** of “Book a table for 2 at 5:30 at Hypothetical Café”. Voice input corresponding to the spoken input is generated by the device **606** and provided to the automated assistant **110**. The automated assistant **110** may convert the voice input to text, and determine that the text indicates an intended action of “restaurant reservation”. The automated assistant **110** may further determine that the text indicates all mandatory parameters for the intended action. Further, the automated assistant **110** may automatically select a particular 3P agent to invoke to perform the intended action with the parameters. For example, the particular 3P agent may be selected based on being the only 3P agent that can perform the intended action or based on user actions and/or input indicating that particular 3P agent as preferred for the intended action. In response to invocation, the particular 3P agent performs the intended action and transmits responsive content to the automated assistant **110** confirming the performance and/or other data. The automated assistant **110** then provides responsive output **482E** that indicates the intended action was performed by the particular 3P agent and additional data provided by the 3P agent (“Email confirmation is on its way”).

In FIG. 5, spoken input **580A** is similar to spoken input **480A** (FIG. 4), but also doesn’t specify the mandatory parameter of “movie name”. In response, the automated assistant provides a prompt **582A** that solicits values for both the “theatre location” and “movie name” parameters.

The user then provides spoken input **580B** that specifies values for both parameters. The automated assistant **110** determines that, of multiple “buy movie ticket” 3P agents, 3P agent 3P1 is the only one that can perform the intended action with the values determined based on the spoken inputs **580A** and **580B**. The automated assistant **110** provides output **582B** indicating this and indicating that the user will be transferred to 3P agent 3P1. In other implementations, output **582B** may not be provided and the user instead directly transferred to 3P agent 3P1.

The automated assistant **110** invokes 3P agent 3P1 with the determined values, and prompt **582C** is provided based on responsive content from 3P agent 3P1. The spoken input **580C** confirms that the user wants to purchase movie tickets from 3P agent 3P1. Output **582D** is then provided based on responsive content from 3P agent 3P1 to confirm the purchase.

FIG. 6 illustrates spoken inputs **680A-F** and responsive outputs **682A-E** and a situation where a third party agent 3P4 is initially invoked and steers the dialog in outputs **682C** and **682D**. However, in response to the spoken input **680E** that indicates a desire to switch to the other capable 3P agent 3P5, the automated assistant **110** intervenes and invokes 3P agent 3P5 while terminating 3P agent 3P4 from the dialog. Output **682E** is then provided based on responsive content from 3P agent 3P5. It is noted that in invoking 3P agent 3P5, the automated assistant **110** provides values for parameters determined based on spoken inputs **680B** and **680D**. Accordingly, the 3P agent 3P5 can generate responsive content that takes into account a value (“large car”) for a parameter (“car size”) determined during dialog steered by 3P agent 3P4. This may lead to an improved user experience and may lessen computational resources (e.g., that would otherwise